Hyperion: A Titan of High-Performance Computing Systems

Hyperion is a large-scale test bed for high-performance computing technologies.

EVER bring home a new computer or software program only to find that it doesn't work as expected? Prerelease testing by the manufacturer can rarely evaluate all of the functions imagined by potential users. In performing full-scale operations, users often uncover hidden errors, or "bugs," that must be corrected in later upgrades.

Scientists working with high-performance computing (HPC) systems can empathize with that frustration. A fundamental challenge they face is establishing a testing phase rigorous enough to find the flaws in hardware or software designed for petaflops computing, where 1 petaflops is 1 quadrillion (10^{15}) floating-point operations per second.

To solve this problem, scientists in Livermore's Computation Directorate have teamed with 10 computing industry leaders to create Hyperion, the world's largest HPC test bed for developing, testing, and scaling Linux cluster technologies. A Linux cluster is a group of thousands of linked computers that together operate as a single, more powerful computing system. Each computer, or node, in the network uses the open-source Linux software as its operating system.

Mark Seager, who leads the Hyperion project at Livermore, is confident the large-scale test bed can find the bugs that often remain undiscovered on smaller testing systems. "Our slogan is, 'If you can make it, we will break it,'" says Seager. Hyperion provides
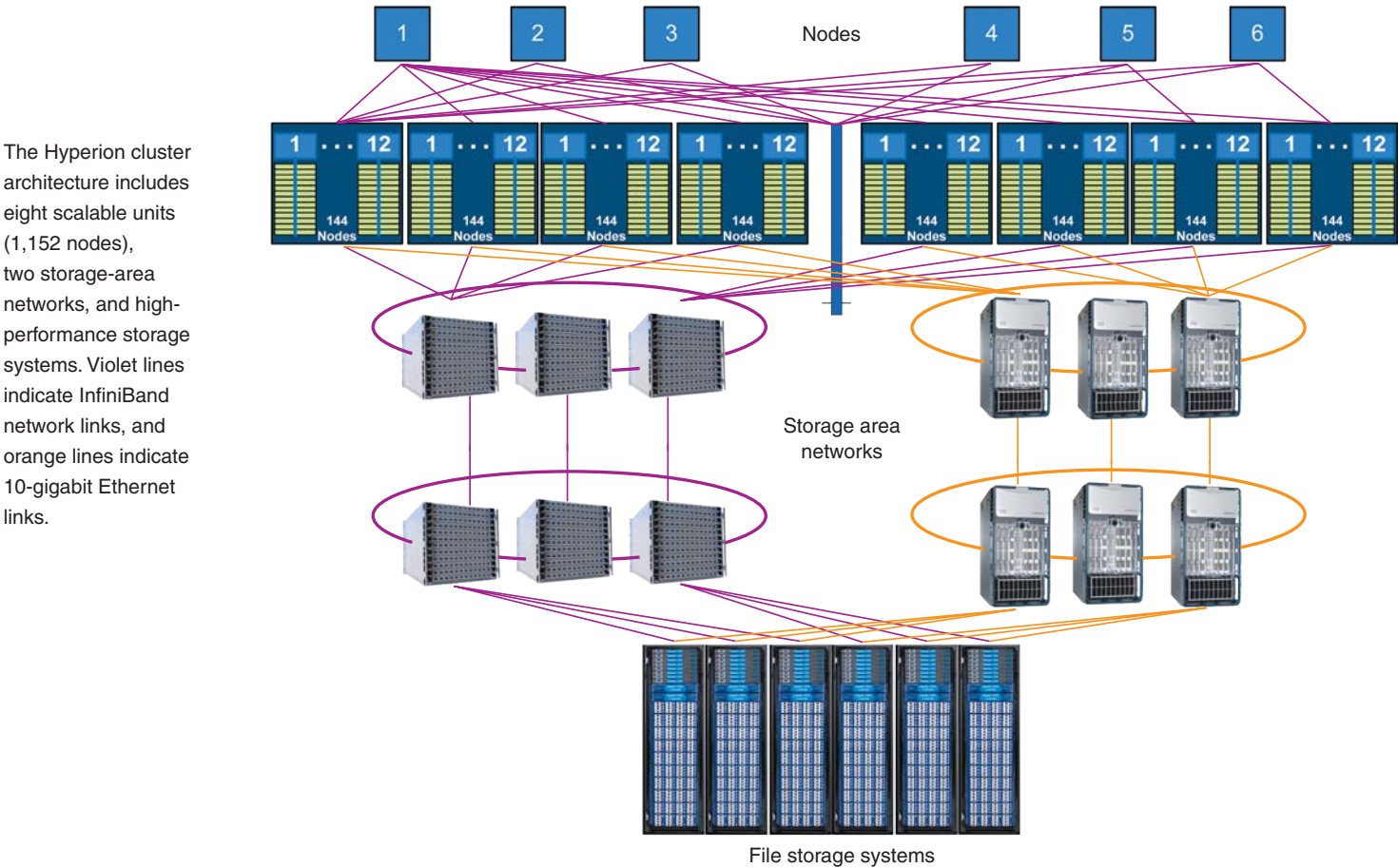
an opportunity for the industrial partners to test their products in a realistic HPC environment, allowing them not only to improve new technologies but also to decrease the time it takes to bring those products to market. The collaboration also promotes long-term relationships between Laboratory researchers and the industrial partners, fostering continuity in HPC technology development.

Michael Dell, the chief executive officer of Dell, Inc., first announced the Hyperion project during his keynote speech at SC08, the annual International Conference for High Performance Computing, Networking, Storage, and Analysis. The National Nuclear Security Administration (NNSA) funds about half of the project as part of the Advanced Simulation and Computing (ASC) Program, which is a key component of NNSA's Stockpile Stewardship Program. The remainder is funded collectively by the 10 industrial partners: Dell, Inc.; Intel Corporation; Super Micro Computer, Inc.; QLogic Corporation; Cisco Systems, Inc.; Mellanox Technologies, Ltd.; DataDirect Networks, Inc.; LSI Corporation; Red Hat, Inc.; and Sun Microsystems, Inc.

The "Hype" in Hyperion

Installed at Livermore, Hyperion is available to the ASC tri-laboratory community—Lawrence Livermore, Los Alamos, and Sandia national laboratories—for developing the HPC technologies needed to maintain the nation's nuclear weapons stockpile without underground nuclear testing. Before deploying large-scale Linux clusters for ASC production applications, Hyperion team members run tests on sets of components as they scale up a new system. Tests are designed to evaluate operating systems, high-performance networking software, and the parallel file systems that distribute data across the servers. In addition, each industrial collaborator is allotted computing time in proportion to its funding contribution to run full-scale system tests of new products before they are released to market.

By rigorously testing a product at scale prior to release, companies can make improvements at the development stage, which is more cost-effective than fixing bugs after a product is deployed. This approach allows the computing industry to make

The Hyperion cluster architecture includes eight scalable units (1,152 nodes), two storage-area networks, and high-performance storage systems. Violet lines indicate InfiniBand network links, and orange lines indicate 10-gigabit Ethernet links.

petaflops technologies more affordable and thus accessible to commerce, industry, and private research and development.

Hyperion is also helping collaborators develop the storage systems and storage-area networks required for ASC's next-generation supercomputer. Called Sequoia, this 20-petaflops system will be delivered to Livermore in 2011. ASC Sequoia will run large suites of complex simulations, allowing scientists to build more accurate models of physical processes, such as those occurring as a nuclear weapon detonates, and to explore frontier science and breakthrough technologies.

Working with a full-scale test bed will establish a blueprint for future petascale computing platforms by helping researchers develop and test processors, memory, networks, storage systems, and visualization technologies. "Hyperion represents a new way to do business," says Seager. "Collectively, we are building a system none of us could have built individually."

One Project, Two Phases

The Hyperion cluster was installed at the Laboratory in two phases. "All of the collaborators were interested in working on the system in 2008," says Matt Leininger, deputy director for Advanced Technology Projects in the Computation Directorate. "However, we wanted some of the hardware to include the new Intel Nehalem processors, which wouldn't be available until early 2009. Splitting the project into two phases was a compromise between starting quickly and waiting for the faster processors."

The first phase, which consisted of 576 nodes, was installed in September 2008. With that cluster, researchers tested the Lustre parallel file system and TOSS, the Red Hat–based operating system that supports several Linux clusters developed by ASC. They also evaluated software used by other HPC researchers, such as the OpenMP message-passing interface and the OpenFabrics high-performance networking software.

In the second phase, which was completed in May 2009, Hyperion doubled in size to 1,152 nodes. The additional nodes incorporated the Nehalem processors and increased the on-node memory by 50 percent, from 8 to 12 gigabytes. As a result, Hyperion has more than 11 terabytes (trillion bytes) of memory—enough to store about 450 high-definition movies—and a peak processing capability of about 90 teraflops.

The Power of 10

Each of the 10 industry leaders plays a vital role in the Hyperion partnership. Dell, Intel, QLogic, Mellanox, and Super Micro Computer built the processors and nodes and helped integrate the input/output system. QLogic, Cisco Systems, and Mellanox built the InfiniBand and Ethernet network components. DataDirect Networks, Sun Microsystems, and LSI created the storage hardware, while Red Hat is responsible for Linux testing and various system administration duties.

The industrial partners collectively contributed about $5.5 million to the Hyperion project, and NNSA contributed approximately $5 million on behalf of Livermore. Fair market value for the system is estimated between $15 and $20 million—a good investment indeed. "We're sharing the cost between NNSA and collaborators to build a place where people can test their equipment, and they don't have to front the full bill for the system," says Lynn Kissel, who recently retired as deputy ASC program leader. "ASC, in some sense, is the glue that makes this partnership happen because we're providing an environment where competitors can share a resource."

In 2009, *Federal Computer Week* selected Seager as one of the Federal 100 top executives from government, industry, and academia who had the greatest impact on government information systems in the past year. Seager credits the collective effort of the Hyperion collaboration, which he says allows the partners to build a scalability test bed that none could afford to build alone.

A Bright Future for Hyperion

"The Hyperion project will advance the state of the art in a cost-effective manner," says Seager. "It offers benefits both to the end users, such as the national security laboratories, and to the computing industry, which can expand the market with proven, easy-to-deploy, large- and small-scale Linux clusters."

Hyperion will help fulfill NNSA goals to provide computing capabilities for national security and to meet the nation's challenges in energy, climate, and other enduring needs. It will also promote scientific discovery in basic science and enhance U.S. competitiveness in HPC. "The Hyperion collaboration will help ensure continuity in developing petascale Linux clusters and the storage technologies for future HPC systems," says Leininger. "As a result, this project will lead to a wide range of economically viable products."

—Kristen Light

Key Words: high-performance computing (HPC), Hyperion, Linux cluster, petascale system.

For further information contact Mark Seager (925) 423-3141 (seager1@llnl.gov).